

	201	211	221	231	241	251	261	271	281	291
	-----G-----	-----G-----	-----G-----	-----G-----	-----G-----	-----G-----	-----G-----	-----G-----	-----G-----	-----G-----
SEQ 3	LAHAGRKATT	VAPW-----	-----ISES	AIATEKVGGW	PDAVKGPDI	P-----	-----FABPFAPK	KA-----	-----MTLDEIE	QFKK-DWVA
SEQ 6	LAHAGRKAST	VAPW-----	-----LSAN	DTASEKMGW	PGRVKGPTNV	P-----	-----FTVKNPVP	KE-----	-----MTKQDIE	DLKT-AWVA
SEQ 8	LAHAGRKASA	VAPW-----	-----LAAQAGKS	LKADESVGW	PADVVGPSGG	E-----	-----EHIF	RA-----	-----LSTAEVR	QVVA-AFAK
SEQ 10	LAHAGRKASN	IAFW-----	-----LQNG	IVATEKVGGW	PORVIGESTV	P-----	-----FHETFEFP	KA-----	-----MTKQDIE	QFKK-DWVA
SEQ 12	LGHGGRKASG	QFLP-----	-----LHLE	QVADKSVNGW	ADKAVAFSAL	A-----	-----FRPNGNLP	VNE-----	-----LTKDEIK	RVVK-DFAA
SEQ 14	LNHAGRKIVE	GVFP-----	-----QOTQHG	QEHCVGSETE	P-----	-----FSOSHWTP	RE-----	-----LTVNEIN	SLVE-DFAA	ELVE-AWKA
SEQ 16	LAHAGRKAST	KAPW-----	-----HYQKGS	ELAGPEQGG	PEHWAPSAI	S-----	-----YHEETFEFP	KE-----	-----MTVLEIN	SLVE-DFAA
SEQ 19	LNHPGRQSM	GAGT-----	-----RGLM	E-KNVAESP	B-----	-----LVLGEAEV	RLSKVLEGT	PRELTVAEIK	DIV-QKFAVT	GLVT-SFVA
SEQ 22	LAHAGRKAST	KAPWHDSETP	SGEYKPREGL	QVVGPEYGGW	PDDVWAPSAI	P-----	-----FSEDFEFP	KE-----	-----MTVEEIE	GLVT-SFVA
SEQ 24	LAHAGRKAST	TAPW-----	-----RG-Y	TVATEAQGGW	ENDVYGFTTN	E-----	-----DRWDENHAQP	HK-----	-----LTKQYD	ELVD-KEVVA
SEQ 27	LAHAGRKAST	TAPW-----	-----ITEARGK	ALAQSENGW	PDDVWAPSAI	P-----	-----YTKDWATP	RE-----	-----LSTE SR	VWVK-KFAES
SEQ 30	LSHAGRKASC	VSPW-----	-----LSVN	AVAAEEVGW	PONIVAPSAI	A-----	-----QENGWNPVP	KA-----	-----FTKEDIE	KLQD-KFVS
SEQ 33	IGHAGRKAST	VVPW-----	-----LDRK	NTAF?	-----	-----	-----	-----	-----	-----
SEQ 35	LSHAGRKASC	VVPW-----	-----LSIN	AVAAKEVGW	PONIVAPSAI	A-----	-----QEAGVNPVP	KA-----	-----FTKEDIE	ELKN-DFLAA
SEQ 40	LAHAGRKAST	VAPW-----	-----LSGG	DVAGEDVNGW	PQDVWAPSAI	P-----	-----WNEKHAVP	KE-----	-----MSLDDIE	AFKK-AFGEA
SEQ 42	-----	-----	-----LPS	KRAGKEAGW	PEDVVGPSGG	EDFTWDERSS	SDPSGGYAP	RE-----	-----LSVREIK	EMVK-DWATA
SEQ 44	VGHPRQARG	SVQ-----	-----QHPISASD	VOLKQEM	-----	-----	-----FGSKGVVP	RP-----	-----ATKEDIK	AVIE-GFAHT
SEQ 83	IGHAGRKASC	VAPW-----	-----LDAG	IAAEKAGGW	PDDVVGPSME	P-----	-----FAPGFTPT	RA-----	-----ITLEEIE	QLKE-DYVSG
SEQ 85	LAHAGRKASD	WSPF-----	-----YRGEKKQ	KFVTQEEGW	PORVWAPSAI	A-----	-----YAGGHVTP	RA-----	-----LTTEDIN	KLQD-KFVS
Bacteria										
T44612	LAHAGRKASA	NPFW-----	-----EGDD	HIGADARGW	-----ETIAPSAI	A-----	-----FGAHLNPV	PRA-----	-----MTLDOIA	RVKQ-DVDA
NP_625402	LAHAGRKAST	KQFW-----	-----RGG	AEVGAADVW	-----QPLASAL	A-----	-----FEDRHPVP	TE-----	-----LTVPQIQ	EAVG-RFADA
NP_295913	LAHAGRKAST	TAPW-----	-----RGK	GAVPAELGGW	-----QVIGEDEN	S-----	-----FHDLEFTP	AM-----	-----MGAEELR	GUVD-AFSAA
AF320254	LGHAGRKAT	KLAW-----	-----EG	IDEPLEAGW	-----ELISASEL	P-----	-----YLPHSQVP	RA-----	-----MTRODMS	RVRN-DEVRA
OYE family										
AF4875	LWATGRADP	DVLA-----	-----DMK-D	LSSS-AVFP	SEKGE-----	-----	-----VGST	LP-----	-----LTEDEIG	QCIA-DEQA
AF4961	LWAGRATIP	QHTG-----	-----SPAVSAS	AVVDSFPEC	YSHPP-----	-----	-----EPVRYADHP	IE-----	-----LTIQ-HL	KQTRDYDCA
Ca2460	LWFLGRVAP	KDL-----	-----DAGLEL	IGPSA-VYV	DESE-----	-----	-----KLAKVGNEL	RE-----	-----LTKEDID	HIVEVEYFNA
Nc4452	LWGLRAANP	EVLA-----	-----KEGGLK	LSSS-AVFP	EEGAP-----	-----	-----VPE	EE-----	-----MTVAEIK	ERVV-EYAAA
ScOYE1	LWVLGWAFF	DTLA-----	-----RDG-LR	YDSASDNYM	NAEQ-----	-----	-----AKKAGANNP	HS-----	-----LTKDEIK	QYIK-EYVQA
ScOYE2	LWVLGWAFF	DTLA-----	-----RDG-LR	YDSASDNYM	NAEQ-----	-----	-----EKAKANNP	HS-----	-----ITKDEIK	QYIK-EYVQA
ScOYE3	LWVLGWAFF	DTLA-----	-----RDG-LR	YDSASDNYM	NAEQ-----	-----	-----EKAKANNP	HS-----	-----LTKDEIK	QYIK-EYVQA
A36990	LWVLGRVANA	KDL-----	-----DSG-LP	LIAPS-AVFP	DENSE-----	-----	-----KLAKVGNEL	RE-----	-----LTKEDID	HIVEVEYFNA
	301	311	321	331	341	351	361	371	381	391
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 3	TKRALIAA-GA	DFVEIHNAHG	YLLSFLSP	-AANRTDQY	G-GSFENRIR	LSLEIAQLTR	DAVGHVP	-----VFLR	ISAS-DWCE	-----ETLPEQ
SEQ 6	VKRAVKA-GA	DFVEIHNAHG	YLLSFLSP	-AVNRTDQY	G-GSFENRIR	LSLEIAQLTR	ENVPKDM	-----VFLR	VSAT-DWLE	-----EVQPNKP
SEQ 8	ARLAVGA-GV	DVIEIHGAHG	YLLSFLSP	-VTNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 10	CKRALIAA-GA	DFVEIHNAHG	YLLSFLSP	-SSNRTDQY	G-GSFENRIR	LSLEIAQLTR	DAVGHVP	-----VFLR	VSAT-DWLE	-----ETLPEQ
SEQ 12	ARRAVEISGF	DAVEIHGAHG	YLLSFLSP	-ISNRTDQY	G-GSFENRIR	FLKEVIDVIR	SSIPNDVP	-----VFLR	ISAS-DWCE	-----DPE
SEQ 14	ARRAVEISGF	DAVEIHGAHG	YLLSFLSP	-ISNRTDQY	G-GSFENRIR	FLKEVIDVIR	SSIPNDVP	-----VFLR	ISAS-DWCE	-----DPE
SEQ 16	AQALKA-GF	DLIEIHGAHG	YLLSFLSP	-ISNRTDQY	G-GSFENRIR	FLKEVIDVIR	SSIPNDVP	-----VFLR	ISAS-DWCE	-----DPE
SEQ 19	ARITAEAG-GF	NGVEIHGAHG	YLLSFLSP	-KTNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	VSAT-DWLE	-----GQPVAAESG
SEQ 22	AKRAVEAG-GF	DIIEIHGAHG	YLLSFLSP	-LSNRTDQY	G-GSFENRIR	FLKEVIDVIR	SSIPNDVP	-----VFLR	ISAS-DWCE	-----DPE
SEQ 24	AKRAVEAG-GF	DIIEIHGAHG	YLLSFLSP	-LSNRTDQY	G-GSFENRIR	FLKEVIDVIR	SSIPNDVP	-----VFLR	ISAS-DWCE	-----DPE
SEQ 27	AKRSNRA-GF	DFVEIHNAHG	YLLSFLSP	-VTNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 30	AKRAIAHA-GF	DFVEIHNAHG	YLLSFLSP	-VTNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 33	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 35	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 38	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 40	AKRAVKA-GV	DVIEIHGAHG	YLLSFLSP	-ITNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 42	AKRAVKA-GV	DVIEIHGAHG	YLLSFLSP	-ITNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 44	AREYLEKA-GF	DGIEIHGAHG	YLLSFLSP	-ATNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 83	ARRAVEAG-GF	DTIDFHGAHG	YLLSFLSP	-ATNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
SEQ 85	ARRAVEAG-GF	DTIDFHGAHG	YLLSFLSP	-ATNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
Bacteria										
T44612	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
NP_625402	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
NP_295913	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
AF320254	TRMPAAG-GF	DIELEHGAHG	YLLSFLSP	-LNNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
OYE family										
AF4875	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
AF4961	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
Ca2460	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
Nc4452	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
ScOYE1	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
ScOYE2	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
ScOYE3	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG
A36990	ARRARDA-GF	EWIEIHGAHG	YLLSFLSP	-HSNRTDQY	G-GSFENRIR	IVREVAARIR	AVIPGMP	-----LFLR	ISAT-DWLE	-----GQPVAAESG

4/17

	601	611	621	631
SEQ 3	GTPIIDPSVY	KQSIIDV		
SEQ 6	AGPYLRKLE	KI		
SEQ 8				
SEQ 10	GTPIIDPKAY	KESIFE		
SEQ 12	QQIVDLIERT	SKLEVN		
SEQ 14				
SEQ 16	PKKLTTPV			
SEQ 19	WIVEKLGMKS	IVGAGVEVTW	YVSELKKLAK	F
SEQ 22	ARI			
SEQ 24	--KLR			
SEQ 27				
SEQ 30				
SEQ 33	KKVVKSSL			
SEQ 35				
SEQ 38				
SEQ 40				
SEQ 42				
SEQ 44	PFDISNADEV	ARVTQLMAEG	KV	
SEQ 83	KKNAFKLVL			
SEQ 85	HRVHVAKK			
Bacteria				
T44612	RYR			
NP_625402				
NP_295913				
AF320254	ETNLQRARRA	VAGK		
OYE family				
Af4875	YLDYPESAAY	MALHNFV		
Af4961	KCYVDYPPAT	ASS		
Ca2460	YNSYDESEKQ	VIGKPLV		
Nc4452	YIDQPSKEP	EKVYGAQA		
ScOYE1	YIDYPTYEEA	LKLGDNRK		
ScOYE2	YIDYPTYEEA	LKLGDNRK		
ScOYE3	YTDYPTYEEA	VDLGNRKN		
A36990	YNSYDESEKQ	VIGKPLA		

Figure 1. A multiple alignment of the 2031 OR amino acid sequence from *A. fumigatus* (SEQ ID No3) along with related 2031 ORs from other fungi and bacteria (see Example 4) and OYEs. Regions 1-11, marked with * or #, refer to amino acids conserved between ORs but not OYEs.

Fungal 2031 ORs are given by the following SEQ ID No.: *A. fumigatus*, SEQ ID Nos. 3, 6 and 8; *A. nidulans*, SEQ ID No. 10; *C. albicans* SEQ ID Nos. 12 and 14; *N. crassa*, SEQ ID Nos. 16 and 19; *M. grisea* SEQ ID Nos. 22 and 44; *S. pombe* SEQ ID No. 24 (NP_595868); *C. trifolii* SEQ ID No. 27; *F. sporotrichioides* SEQ ID Nos. 30, 33 and 35; *F. graminearum* SEQ ID Nos. 38 and 83; *M. graminicola* SEQ ID Nos. 40 and 42; *U. maydis* SEQ ID No 85.

Bacterial ORs resembling 2031 are:

T44612 (*Pseudomonas putida*), SEQ ID No. 86; NP_625402 (*Streptomyces coelicolor*), SEQ ID No. 87; NP_295913 (*Deinococcus radiodurans*), SEQ ID No. 88; AF320254 (*Azoarcus evansii*, SEQ ID No. 89.

Fungal ORs similar to the Old Yellow Enzyme family (originally identified in *S. cerevisiae*):

A. fumigatus, Af4875 and Af4961, SEQ ID Nos. 90 and 91 respectively; *C. albicans*, Ca2460 and A36990, SEQ ID Nos. 92 and 93 respectively; *N. crassa*, Nc4452, SEQ ID No. 94; *S. cerevisiae*, OYE1, OYE2 and OYE3, SEQ ID Nos. 95-97 respectively.

Details of the sequence searches that identified the ORs other than SEQ ID No. 3, and methods for the construction of multiple alignments are given in Example 4 hereinafter.

6/17

	301	311	321	331	341	351	361	371	381	391	
SEQ 1	GACTGTGCGC	GATATCGACG	TTCTCTCTGC	CGAGGGGATC	CCCTACTTCA	CTCCGGGCGCA	GAACCTCTCT	GCCGGTACGG	CAGCTAACCC	CCAGACCAAT	
SEQ 2	GACTGTGCGC	GATATCGACG	TTCTCTCTGC	CGAGGGGATC	CCCTACTTCA	CTCCGGGCGCA	GAACCTCTCT	GCCGGTACGG	CAGCTAACCC	CCAGACCAAT	
SEQ 4	TGTTGTGCGT	GACATCGAGA	ACAAACCCGC	GCCGGGTATC	TCGTACTTTA	CTCCGGGCGCA	AGAGCCGCGT	GCTGGCACCG	CTGCTAATCC	TCAGTCTGAT	
SEQ 5	TGTTGTGCGT	GACATCGAGA	ACAAACCCGC	GCCGGGTATC	TCGTACTTTA	CTCCGGGCGCA	AGAGCCGCGT	GCTGGCACCG	CTGCTAATCC	TCAGTCTGAT	
SEQ 7	CGCTTCTCGG	TCCCGCGCGG	TCACCAAGTC	CTCTCTCACC	CCCTACTTCA	CTCCGGGCGCA	CAATGGAGGC	GCCGCCCTGC	ACCCCGACGA	CCCCAC----	
SEQ 9	GGCTCTCTCT	GACGTGAAAA	ACACCCCGCG	CGCCGGGATC	CCCTACTTTA	CACGAGCACA	GAACCTCTCT	GCTGGAACAG	CTGCCAACCC	GCAACCCAGC	
SEQ 11	AGTAAACCA	TCAGATGAAA	TCAAAGGTGC	TCCTGAGGTT	TCCTATTACA	CTCCAGAACCA	GCCTGTTCGG	GCTGGTACTT	TTTATCCCCA	ATCGTC----A	
SEQ 13								---ATGAAA	ACAACAATAC	TATACCG---	
SEQ 15	CACCCAGAA	AAGACTCTCT	CCCCCGCGGC	CGCCGGGTGT	CCCTTCTACA	CCCCGGGCGCA	GGTCCCCGCG	GCCGGCACTC	CCCTCCCTCC	CACCCCT---	
SEQ 17								ATGGCTACTT	CCACTACCTC	CGACCTC---	
SEQ 18								ATGGCTACTT	CCACTACCTC	CGACCTC---	
SEQ 20	GGCAGAAAA	AAGACTTTGA	GCAAAACCGGC	CGCCGGGGTG	CCTTACTACA	CCCCAGCCCCA	GGAGCCGCGG	GCAGGACCCG	CTTTGCAGCA	GCAGGACG--	
SEQ 21	GGCAGAAAA	AAGACTTTGA	GCAAAACCGGC	CGCCGGGGTG	CCTTACTACA	CCCCAGCCCCA	GGAGCCGCGG	GCAGGACCCG	CTTTGCAGCA	GCAGGACG--	
SEQ 23	---ATGAC	TATTGTTAAT	GAAGGAGCCG	AAATGTTGG	TTATTTTACA	CCTGCGGCAAA	AAATACCAGC	TGGAGCGCGG	ATAGGTGTAC	CGCAAA----	
SEQ 25	CAGCATGACG	GGCAGCGCGA	ACAAGGCCGC	CGCCGGGTGT	CCCTTTTACA	CCCCGGGCGCA	GGAGCCCTCC	CGGGAACGCG	CAGCTGCAGC	CAGCAGCG--	
SEQ 26	---ATGACG	GGCAGCGCGA	ACAAGGCCGC	CGCCGGGTGT	CCCTTTTACA	CCCCGGGCGCA	GGAGCCCTCC	CGGGAACGCG	CAGCTGCAGC	CAGCAGCG--	
SEQ 28	GGCTTACGAG	ATAATCGACA	ACGTTGCGGC	TGAAGGGGTT	CCATATTACA	CACCGGCTCA	AGACCCGCGA	GCTGGTACGG	AGACAAGCGG	CTCAACG---	
SEQ 29	GGCTTACGAG	ATAATCGACA	ACGTTGCGGC	TGAAGGGGTT	CCATATTACA	CACCGGCTCA	AGACCCGCGA	GCTGGTACGG	AGACAAGCGG	CTCAACG---	
SEQ 32											
SEQ 34	CCATCACAAC	ATCATGATCA	ATAAGGAAGC	TCCGAATGTT	CCTTCTATA	CTCCAGTGCA	AGATCCACCA	GCAGGAACGT	CTTACGATGT	TCAGCTGAA	
SEQ 36	---GCACGAGG	ATTATTGACA	ACATCGCGGC	TGAAGGGGTT	CCCTACTACA	CGCCTGTCTA	AGACYCTCCA	GCAGGCACAC	AGACCAGCGG	CTCAACG---	
SEQ 37	---GCACGAGG	ATTATTGACA	ACATCGCGGC	TGAAGGGGTT	CCCTACTACA	CGCCTGTCTA	AGACYCTCCA	GCAGGCACAC	AGACCAGCGG	CTCAACG---	
SEQ 39											
SEQ 41											
SEQ 43					---ATGT	CCCCACCAGG	CTTCGAAGCG	GCCCTGCGCG	ACCCCTCACC	GCTCGGC---	
SEQ 42	AAACAAGGAG	GTGTTTCAGA	ATGTCGCTGC	CAAAAGGAGT	CAATCTGTCA	ACCTGAGACA	AGACTGCTCA	CCAGGCTCTG	GTATAACCGG	TCCCAAT---	
SEQ 84	ACCCGCTCTC	GTCGACTCGA	TGATGCACT	CAAGATCAGC	AACCTTGTC	CACTCGAAG	TGGCCACCTT	CCTCTGGTCT	CGGTCGCGG	ATCCATCCTG	
SEQ 1	GG-----CC	AGAAGATCCC	CAAGCTCTTC	ACGCCCTTGA	CCATCCGTTG	CGTCACCC--			---TTCCAGAAC	CGCCTTGTTG	
SEQ 2	GG-----CC	AGAAGATCCC	CAAGCTCTTC	ACGCCCTTGA	CCATCCGTTG	CGTCACCC--			---TTCCAGAAC	CGCCTTGTTG	
SEQ 4	GG-----AT	CGGCACCTCC	CAAGCTCTTC	CGGCCGCTTT	CGGTGCGGGG	TCTGACCC--			---TTTCACAA	CGCATTTGGG	
SEQ 5	GG-----AT	CGGCACCTCC	CAAGCTCTTC	CGGCCGCTTT	CGGTGCGGGG	TCTGACCC--			---TTTCACAA	CGCATTTGGG	
SEQ 7	GG-----GACCC	TACGCTCTTC	CGGCCCTTAC	AAATCCGCAA	TGTGACG--				---CTCAAGAAC	CGCATCTG--	
SEQ 9	GG-----CA	ATGCCGTCCC	CAAGCTGTAC	ACACCTCTGA	CGGTGCGTGG	GGTGACCC--			---TTCCACAA	AGACTTGGG--	
SEQ 11	GA-----TG	AAGTTGCTCC	CAAAATTTT	CAACCTTTAA	AGATTGGTAA	GCTTGCT--			---TTGCCAAAC	AGAATTGGG--	
SEQ 13			---GCATTAT	CAACCTATA	AGATCAATGA	CTGATC--			---AC	ATTACCTAAT	
SEQ 15		---G	GCGATGCTCC	TACTCTCTTC	ACCCCTCTCA	AGATCCGTTG	TGTTGAG--			---CTCCAGAAC	CGCTTGCGG--
SEQ 17		---AACTCTCC	CAACCCCTCA	CCCTCCCTCA	TGCGCTT--				---AC	CTCCCTCAAC	
SEQ 18		---AACTCTCC	CAACCCCTCA	CCCTCCCTCA	TGCGCTT--				---AC	CTCCCTCAAC	
SEQ 20		---CAATCCC	AAGCTGTGTC	AAGCTCTGGA	AGATCCGTTG	CGTCGAG--			---CTCTCCAA	CGCTTTGGG--	
SEQ 21		---CAATCCC	AAGCTGTGTC	AAGCTCTGGA	AGATCCGTTG	CGTCGAG--			---CTCTCCAA	CGCTTTGGG--	
SEQ 23		---C	AAATTTATTT	ACTCCTCTTA	AAATTAGAGG	AGTGAGG--			---TTCCATAAC	AGATGTTT--	
SEQ 25		---CTCC	GACGCTCTTC	AAGCCCTCTC	GCATCCGCGA	CCTCACC--			---ATCACAAAC	CGCATCTGG--	
SEQ 26		---CTCC	GACGCTCTTC	AAGCCCTCTC	GCATCCGCGA	CCTCACC--			---ATCACAAAC	CGCATCTGG--	
SEQ 28		---AAGCTATC	ACACCCATCA	CCATCCGCGG	CGTCACA--				---TTCCCAAAC	CGCCTCTTC--	
SEQ 29		---AAGCTATC	ACACCCATCA	CCATCCGCGG	CGTCACA--				---TTCCCAAAC	CGCCTCTTC--	
SEQ 32											
SEQ 34	GG-----	AAGCCTATTC	TCTCTTATTA	AAATAAGAAA	CCTGACT--				---CTTCAAAAC	CGGATTTT--	
SEQ 36		---AGGTTTT	ACACBCATCA	CCATCCGAGG	CGTCACA--				---TTCCCAAAC	CGCTCTCTT--	
SEQ 37		---AGGTTTT	ACACBCATCA	CCATCCGAGG	CGTCACA--				---TTCCCAAAC	CGCTCTCTT--	
SEQ 39			---CCTCA	AGATCCGAGG	TCTTACC--				---CTCCAGAAC	CGTATTATG--	
SEQ 41											
SEQ 43		---ACGC	CGCTCAATA	CCCCGCTCG	GGCGGCTG--				---GGGCCAAAC	CGGTTCTCT--	
SEQ 82	---A	ATACTCTACC	AAAGGTCTTT	ACACCCATCA	AGATTGCGGG	CATGACC--			---ATGCCCAAC	CGTATCTGG--	
SEQ 84	CCAGAGGGTG	TCAAAAAC	GGCTTTGTT	CAACGTTGA	CATTGCCCTT	TGCTGCACCG	GAACAGGCGG	GTAAGATGAC	CTTCAAGAAC	CGCATCATT--	
SEQ 1	TAAGTCCGTT	TGCGCTTGCT	CATATCGACG	AAAGCTAATC	CCCCGTCAG--				---CTCGC	GCCCCCTGCG	
SEQ 2									---CTCGC	GCCCCCTGCG	
SEQ 4	TGAGTGCAGT	CCAGGCAATT	ATGCTATCCA	TCCTATGCGA	GCCCTTGCGAT	TGGAACAGCC	GCTTACAGGG	AATGATTAATG	AGTAGTATC	GCCACTCTGC	
SEQ 5									---CTATC	GCCACTCTGC	
SEQ 7									---GTGTC	GCCCATGTGC	
SEQ 9									---CTCGC	GCCGCTCTGC	
SEQ 11									---GTATC	TCCAATGTGT	
SEQ 13									---GTTTC	ACCAATGTGC	
SEQ 15									---GTTTC	ACCAATGTGC	
SEQ 17									---AAAGC	CGCCATGGCC	
SEQ 18									---AAAGC	CGCCATGGCC	
SEQ 20									---GTCTC	GCCCATGTGC	
SEQ 21									---GTCTC	GCCCATGTGC	
SEQ 23									---GTTTC	GCCCATGTGC	
SEQ 25									---GTGAG	CCCCATGTGC	
SEQ 26									---GTGAG	CCCCATGTGC	
SEQ 28									---CTTGC	CCCTCTCTGC	
SEQ 29									---CTTGC	CCCTCTCTGC	
SEQ 32											
SEQ 34									---GTCTC	CCCAATGTGT	
SEQ 36									---CTTGC	CCCTCTCTGT	
SEQ 37									---CTTGC	CCCTCTCTGT	
SEQ 39									---TTGAG	GGGGCTCTGC	
SEQ 41											
SEQ 43									---AAGCC	GGCCATGTGC	
SEQ 82									---GTGAG	CCCCATGTGC	
SEQ 84									---GTCTC	TCCCATGTGC	

	1501	1511	1521	1531	1541	1551	1561	1571	1581	1591
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 1		GCAGAGCTGG	AAGTCGGAGG	ATACCGTGGC	GTTTCGCCAG	GAGCTGGTCA	AGCAGGGCGC	CGTTGATCTG	ATCGATATCA	GCAGCGGTGG
SEQ 2		GCAGAGCTGG	AAGTCGGAGG	ATACCGTGGC	GTTTCGCCAG	GAGCTGGTCA	AGCAGGGCGC	CGTTGATCTG	ATCGATATCA	GCAGCGGTGG
SEQ 4	---CAA	GCCCACTGG	CGAGGCGTGG	ACACTGTCCG	ATTTGCGAAG	ATCCTGGCAG	AAACGGGTTA	CGTTGACGTG	CTTGACGTGA	GCAGTGGCGG
SEQ 5	---CAA	GCCCACTGG	CGAGGCGTGG	ACACTGTCCG	ATTTGCGAAG	ATCCTGGCAG	AAACGGGTTA	CGTTGACGTG	CTTGACGTGA	GCAGTGGCGG
SEQ 7	---CGCGGAGTC	GGGCAGCTGG	GATAT---GC	AGAGCTCGCT	GGAGCTGGTC	AAGAAGCTGC	CCGAATGGGG	CATTGACCTG	GTGGATGTCA	GCTCCGCGCG
SEQ 9		GGAAATCGTGG	AAGCTCTCTG	ACTCCGTCGG	CTTCGCCGAA	CGCCTCGCTG	CCCAGGGCGC	TATTGACCTG	ATCGACGTCT	CTTCCGGCGG
SEQ 11		---GAAGCTTGG	ACTATTGAAG	ATTTCCAAAA	---AATTAGCT	GACATTTTAG	TAGAAAAGGG	TATTGCTTTG	TTGTGATGTT	CATCTGGTGG
SEQ 13		---GAAGCTTGG	TCTACGGAA	ATGCATTGA	---AGTTGGCC	GACTTTGTTA	TTGATTAGG	AGTAAAGGTG	ATCGACGTTA	CATCAGGTGG
SEQ 15	---GGCCA	GCCTCTGGTG	GACCTCCAGC	AGACCATTTG	---AGCTCGCC	AAGATCCTCC	CCGACCTCGG	CGTCGACCTC	CTCGACGTCT	CTTCCGGCGG
SEQ 17	AGGAGGAGGA	GGAGACGGAT	ACGGCGGAGG	AGGTGTTGA	---AGCAGATT	GAGCTTTTTG	AGCAGTGGGG	GATCGACTTT	GTCCAGGTTA	CGGGTGGCAG
SEQ 18	---GAGGAGGA	GGAGACGGAT	ACGGCGGAGG	AGGTGTTGA	---AGCAGATT	GAGCTTTTTG	AGCAGTGGGG	GATCGACTTT	GTCCAGGTTA	CGGGTGGCAG
SEQ 20	---GA	GCCTAGCTGG	GACCTCGAGC	AGAGCACAC	---AGCTTGCC	AAGCTCCTCC	CGGACCTGGG	TGTCGACCTG	CTCGACGTCA	GCTCGGGCGG
SEQ 21	---GA	GCCTAGCTGG	GACCTCGAGC	AGAGCACAC	---AGCTTGCC	AAGCTCCTCC	CGGACCTGGG	TGTCGACCTG	CTCGACGTCA	GCTCGGGCGG
SEQ 23		---GGATGG	GAGATAGAAG	ATACAGTTG	---CATTAGCA	GCGAGGCTTC	GCGATGGTGG	TGTTGACTTG	ATAGATGTTA	GCTCTGGTGG
SEQ 25		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	ACGAGTTTCC	TGAAAGCTGG	ACAGTCGAGC	AGACTT---G	TCAACTCGCG	CGTATCTTGC	CCAAGCATGG	AGTAGACTTG	GTGGACGTCA	GCTCAGGCGG
SEQ 34		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39		-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	---AAGAAGTT	CGGAAGCTGG	GATGTCGAAA	GCACGATCA	---AGATCTCC	AAAATCCTGG	CCGACTTGGG	CGTTGATCTC	CTCGACGTGT	CTTCCGGTGG
SEQ 43		---GTTTTCAAG	CCA---GAGG	AGGGCGGTGC	---AGTTGTGC	GAGGGCCCTCG	AGGGCCGCGG	CATGGATTTT	GTCCGAGACA	CGGGCGGCAG
SEQ 82	---TACGAGGG	AGAGAGCTTG	ACTCTTGAGG	AGAGCATCA	---AGCTTGCA	CACCGAGTTAG	CAGACGGTGG	TGTCGAGTGT	TTGAGTGTTT	CCAGTGGTCC
SEQ 84	---GC	CGACTCTTGG	ACCTTGTAACC	AGACGGTTG	---AACTCGCC	AAGATGGCTCC	AAGAGCGTCTG	AGTCGACCTG	CTAGACGTCA	GCTCCGGCGG
	1601	1611	1621	1631	1641	1651	1661	1671	1681	1691
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 1	#####	#####	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 2	TGTTCTCGCG	CAG-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	TGTTCTCGCG	CAG-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	CACCTATTTCG	GAG-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	GAACCAACAAG	GAC-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	TGTCACAGCC	GCG-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	TAACGATTAT	AGA-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	AAATGTTTGGC	CAT-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	CAACAACAAG	GAC-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	TTATGAGGAT	CCTCAGGTAA	GTTTTGGTGT	TGTTTGAGGG	ATGGGGCAAG	GGGTTGTCTG	TCGTGAACAA	AAAAGGGGG	ACGGAAACAA	TGCTAACGCC
SEQ 18	TTATGAGGAT	CCTCAG-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	AAACTCGGTG	GCC-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	AAACTCGGTG	GCC-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	TAATCACAAG	GAT-----	-----	-----						

11/17

	1801	1811	1821	1831	1841	1851	1861	1871	1881	1891
SEQ 1	-----AAGCT	GCTGGTTGCC	GCCGTGGGTG	CCATCACC	-----	-----	-----	-----	-----AACG	GCAAGCAGGC
SEQ 2	-----AAGCT	GCTGGTTGCC	GCCGTGGGTG	CCATCACC	-----	-----	-----	-----	-----AACG	GCAAGCAGGC
SEQ 4	-----AAACT	CGCAGTGGCA	TCAGTGGGTA	TGATTGCC	-----	-----	-----	-----	-----AGCG	CGCATTGGCC
SEQ 5	-----AAACT	CGCAGTGGCA	TCAGTGGGTA	TGATTGCC	-----	-----	-----	-----	-----AGCG	CGCATTGGCC
SEQ 7	-----GCTGG	CGCGTCGACT	CTGTGGGTG	CTGTAGGTCT	GATCACCAGT	TCGGAACAGG	CGAGGGGACT	AGTTCAGGGA	CGCGACGAGG	CGACTGCAGC
SEQ 9	-----AAGCT	CCTTGTGGCG	ACGGTGGGCA	CGATCAGC	-----	-----	-----	-----	-----AAGC	GTAAGCAGGC
SEQ 11	-----AAGTT	ATTGGTCAGT	TGCGTTGGTG	GGCTTGAA	-----	-----	-----	-----	-----A	AAGATCCTGA
SEQ 13	-----CGATG	TTTGATCGCA	TGCAGTGGAG	GATTAGAT	-----	-----	-----	-----	-----C	GAGACATATT
SEQ 15	GCAAGCAGCT	CCTCGTCGGT	GCCGTCCGCT	TGGTCACC	-----TCG	GCTGAGATCG	CCAAGGAGAC	CGTGCAGGAG	AAGGAGGATG	GCAGAGTCAC
SEQ 17	TCCCAAGCT	TCCTCTCATG	GTACCGGGCG	GCTTCCGC	-----	-----	-----	-----	-----ACTC	GTCAGGGCAT
SEQ 19	TCCCAAGCT	TCCTCTCATG	GTACCGGGCG	GCTTCCGC	-----	-----	-----	-----	-----ACTC	GTCAGGGCAT
SEQ 20	-----AGGTT	GCTCATAGGC	GCGGTCGGCA	ACATCAAC	-----	-----	-----	-----	-----ACGG	CTGACATTGC
SEQ 21	-----AGGTT	GCTCATAGGC	GCGGTCGGCA	ACATCAAC	-----	-----	-----	-----	-----ACGG	CTGACATTGC
SEQ 23	-----AT	ACTACTGGCG	GCTGTCGGAA	TGATCAGG	-----	-----	-----	-----	-----GATG	GTCTTACGGC
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----AGTGT	ACTTGTTC	GCAGTAGGTG	GAATCAAG	-----	-----	-----	-----	-----A	CTGGACATCT
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----ATGGT	GGTCTACACC	ACCGGCGGCT	TCAAGACG	-----	-----	-----	-----	-----GTGGGC	CCATGGTCCGA
SEQ 45	-----AAGAT	GTGATCAGC	ACTGTTGGTA	GCATCAAG	-----	-----	-----	-----	-----ATAG	GTACCTTTGC
SEQ 47	-----ATCGAAC	CGACGCGTCC	AAACGCATGC	TCGTCCGGG	-----	-----	-----	-----	-----CCGTGG	GAATGATGGA

	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991
SEQ 1	GAATCAG	-----ATTCTAG	AGGAGCAG	-----	-----	-----	-----	-----	-----	-----
SEQ 2	GAATCAG	-----ATTCTAG	AGGAGCAG	-----	-----	-----	-----	-----	-----	-----
SEQ 4	CAATTCC	-----TTGTTGG	AGAAGGAC	-----	-----	-----	-----	-----	-----	-----
SEQ 5	CAATTCC	-----TTGTTGG	AGAAGGAC	-----	-----	-----	-----	-----	-----	-----
SEQ 7	CGAGGCAATG	CTGTCCGGAC	CTGAACCC	-----	-----	-----	-----	-----	-----	-----
SEQ 9	GAACAAG	-----CTGCTTG	AGGAGGAG	-----	-----	-----	-----	-----	-----	-----
SEQ 11	ATTGCTCAAC	AAATATTTAG	AAGAAGGA	-----	-----	-----	-----	-----	-----	-----
SEQ 13	TAAACTCGAT	GAGTTTATTG	CTAATGGT	-----	-----	-----	-----	-----	-----	-----
SEQ 15	CATCCAGCGC	GAGAACGGCG	CCAAGACT	-----	-----	-----	-----	-----	-----	-----
SEQ 17	GGAGGCC	-----GCTTTGG	AATCCGAT	-----	-----	-----	-----	-----	-----	-----
SEQ 18	GGAGGCC	-----GCTTTGG	AATCCGAT	-----	-----	-----	-----	-----	-----	-----
SEQ 20	GCGGATGTC	GTGGATGAGC	AGGGCGCCGA	GAAGGTGGCC	GAGGCCAAGC	AGACGCATGA	CACCATCGAG	GTCGTGAGCG	AATCAGATGG	CGGCAAGACC
SEQ 21	GCGGATGTC	GTGGATGAGC	AGGGCGCCGA	GAAGGTGGCC	GAGGCCAAGC	AGACGCATGA	CACCATCGAG	GTCGTGAGCG	AATCAGATGG	CGGCAAGACC
SEQ 23	GAATGAAATC	CTAGAAAGTG	GAAGAGCT	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	TGCTGAA	-----GAGGTTT	TGCAATCT	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	CGCGCTGCAG	GGCGTCGATG	GG-----	-----	-----	-----	-----	-----	-----	-----
SEQ 45	GGAGGAG	-----ATCATCG	CTGGAGGAGA	GGACGATACC	-----	-----	-----	-----	-----	-----
SEQ 47	AGGTTC	-----TACGATT	CGCCCAAC	-----	-----	-----	-----	-----	-----	-----

	2001	2011	2021	2031	2041	2051	2061	2071	2081	2091
SEQ 1	GATATCGACG	TTGCGCTGGT	TGGCCGTGGG	TTCCAGAAGG	ATCCCGGTCT	GGCCTGGACG	TTTGCTCAGC	ACCTCGGCGT	C-----	-----
SEQ 2	GATATCGACG	TTGCGCTGGT	TGGCCGTGGG	TTCCAGAAGG	ATCCCGGTCT	GGCCTGGACG	TTTGCTCAGC	ACCTCGGCGT	C-----	-----
SEQ 4	GGACTGGACC	TTGTGCTGGT	TGGAGTGGGC	TTCCAGAAGA	ACCCGGGGCT	GGTGTGGGCG	TGGGCGGACG	AGCTGAATGT	A-----	-----
SEQ 5	GGACTGGACC	TTGTGCTGGT	TGGAGTGGGC	TTCCAGAAGA	ACCCGGGGCT	GGTGTGGGCG	TGGGCGGACG	AGCTGAATGT	A-----	-----
SEQ 7	AAGGCGGATG	CCATTCTGAT	AGCCCGTCAG	TTCTTGGCGG	AGCCAGAATG	GGTGTTTTCC	ACGGCGAGAA	AGTTGGGCGT	G-----	-----
SEQ 9	GGATTGGATG	TTGCGCTTGT	GGGACGTGGT	TTCCAGAAGG	ATCCCGGTCT	GGCCTGGACT	TTGCGGACG	ATCTTGATGT	T-----	-----
SEQ 11	ACATTGTGATC	TTGCTTTGAT	CGGTAGAGGA	TTTTTAAGAA	ATCCAGGTTT	GGTATGGGAG	TTTGCCGATA	AACCTGGTGT	T-----	-----
SEQ 13	GACTTTTGATA	TAGCATTGAT	AGGTAAGGA	TTTCTCAAAA	ACACTGGATT	GATCAGCCGT	ATTGCTGACC	AATTGCAAGC	A-----	-----
SEQ 15	CGTGCCGATA	TGGTCCTTGT	TGCCAGGCAG	TTCTTAAGAG	AGCCCGAGTT	CGTCTCTACT	GTGCGCGACG	AGTTGGGTGT	T-----	-----
SEQ 17	GATTGCGACA	TGATCCGAT	CGGACGCCCG	GCCATCATCA	ACCCTTCGCT	TCCCGCCCAAC	TTGATCTCTCA	ACCCGAGGTT	G-----	-----
SEQ 18	GATTGCGACA	TGATCCGAT	CGGACGCCCG	GCCATCATCA	ACCCTTCGCT	TCCCGCCCAAC	TTGATCTCTCA	ACCCGAGGTT	G-----	-----
SEQ 20	AAGGCGGATC	TGGTCCCTCAT	TGCTCGCCAG	TTCTTGGCGG	AGCCTGAGTT	TGTGCTGAGG	ACGGCGCATA	ACCTTGGGGT	C-----	-----
SEQ 21	AAGGCGGATC	TGGTCCCTCAT	TGCTCGCCAG	TTCTTGGCGG	AGCCTGAGTT	TGTGCTGAGG	ACGGCGCATA	ACCTTGGGGT	C-----	-----
SEQ 23	-----GATG	TTACTTTTGT	CGCAAGGGAG	TTCTTAAGGA	ACCCGTCTGT	GGTGCTAGAC	AGCGCGAACC	AGTTGGGTGA	A-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	GGTATCGACA	TTGTGAGGGC	TGGACGTTGG	TTCCAACAGA	ATCCTGGTCT	GGTTCGAGCT	TTTGCTAACG	AGCTTGGCGT	G-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----ATAGGCAT	CGGGCGCGCA	GCCGGTTCGG	AGCCGGACCT	CGCCAAGGAC	ATCATCGCGG	GCAAGGTGTC	CAGCATTATC	AAATACGCCA
SEQ 45	CCCTTGGATC	TTGTGGCTTC	AGGCCGCTCG	TTCCAGAAGA	ACACTGGACT	TGTTTGGTCA	TGGGCTGACG	ATCTGAACAC	T-----	-----
SEQ 47	GGCCAAGACC	GCAGCCAGAT	TGGCAAGTTG	GCCGAGCAGT	CGATTTCAGAG	CGGAGAGTGT	GATGCGGTAC	TGTTGGCACG	T-----	-----GGATTGA

13/17

	2401	2411	2421	2431	2441	2451	2461	2471	2481	2491
SEQ 1	CGTCCTCTTA	AGTTTCTCCG	TCATTCGTTC	TATTCTACTC	CAATCGCAAC	GCATGGCGAC	CACGGATCGA	GTCGAATTTC	TCCGTCGTTC	GTATCTGATC
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

	2501	2511	2521	2531	2541	2551	2561	2571	2581	2591
SEQ 1	AATATAAAAA	GCGGGGAATG	GCTTGACCCC	GCGCAGAATG	TCGATCTCTT	CGCAAACCTCT	CGGTGTATAG	GACGCTCAGC	AACGATCAAG	G
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Figure 2. A multiple alignments of the 2031 OR nucleic acid sequence from *A. fumigatus* (SEQ 1,2) along with related 2031 ORs from other fungi and bacteria (see also Example 4). Regions 1-11, marked with * or #, refer to regions conserved at the amino acid level between Ors but not OYEs.

Fungal 2031 ORs are given by SEQ ID No.: SEQ ID Nos. 1, 2, 4, 5, and 7, *A. fumigatus*; SEQ ID No. 9, *A. nidulans*; SEQ ID Nos. 11 and 13, *C. albicans*; SEQ ID Nos. 15, 17 and 18, *N. crassa*; SEQ ID Nos. 20, 21 and 43, *M. grisea*; SEQ ID No. 23 (NP_595868), *S. pombe*; SEQ ID Nos. 25 and 26, *C. trifolii*; SEQ ID Nos. 28, 29, 31, 32 and 34, *F. sporotrichioides*; SEQ ID Nos. 36, 37 and 82, *F. graminearum*; SEQ ID Nos. 39 and 41, *M. graminicola*; SEQ ID No. 84, *U. maydis*.

14/17

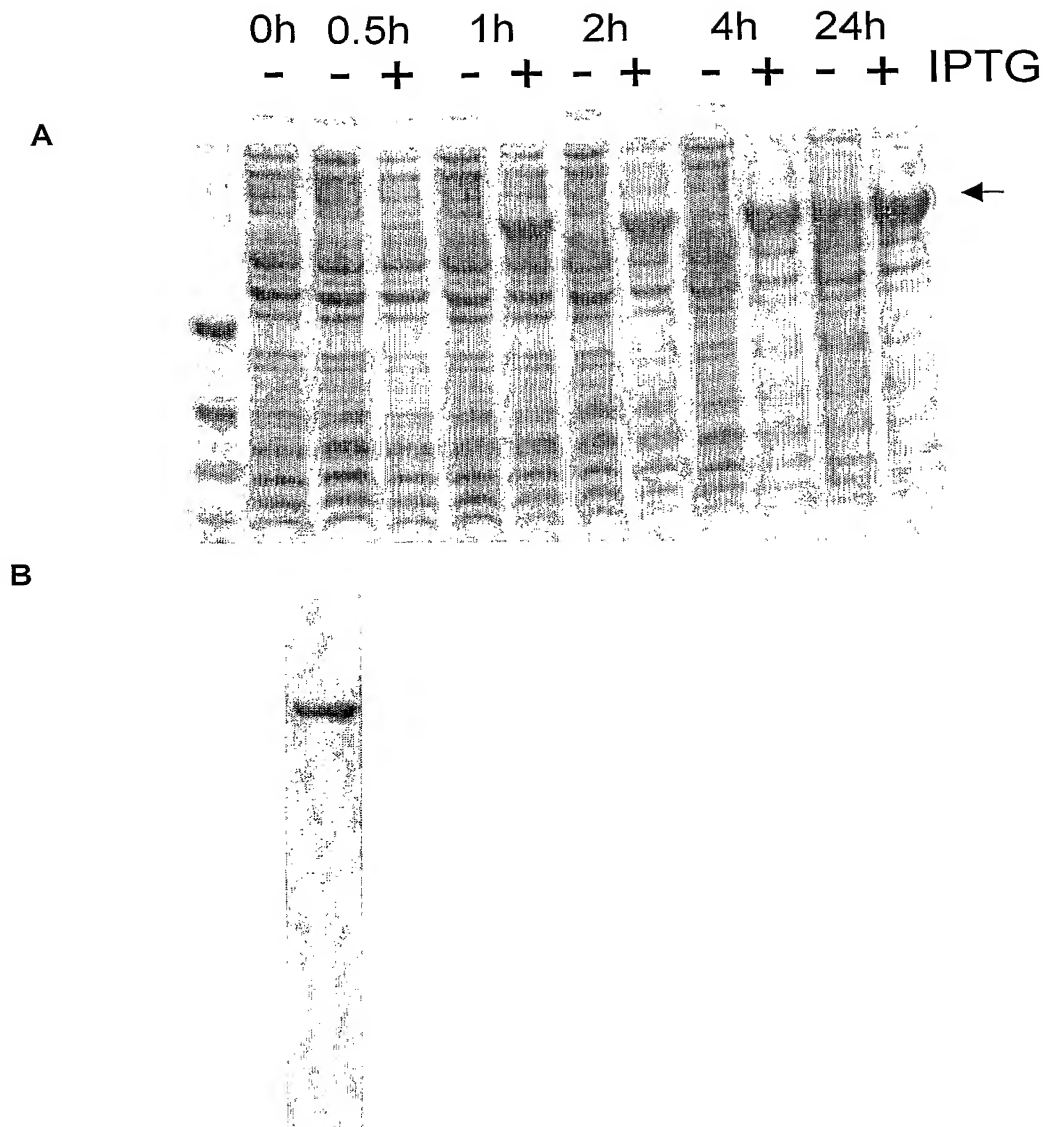


Figure 3. Recombinant 2031 OR. (A) Time course of recombinant 2031 OR induction over 24 hours after the addition of IPTG (samples without IPTG are also shown). The gel was stained with coomassie; A prominent band of the correct molecular weight (marked with an arrow) is seen. (B) Coomassie stained gel showing purified recombinant 2031.

15/17

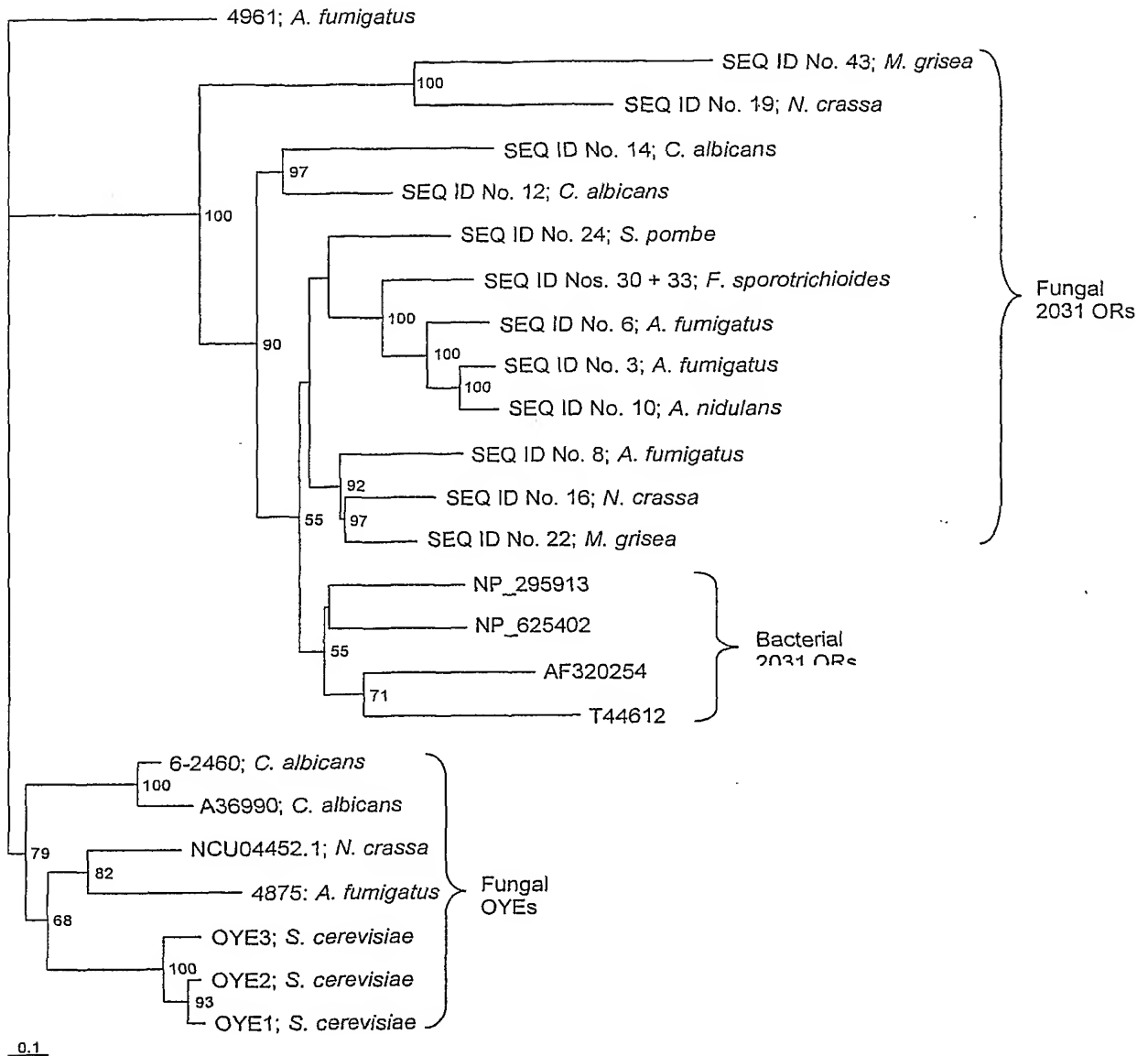


Figure 4. Phylogenetic tree showing relationships between *A. fumigatus* 2031 OR and similar proteins. This demonstrates a 2031 OR clade, which can be distinguished from the OYE proteins.

16/17

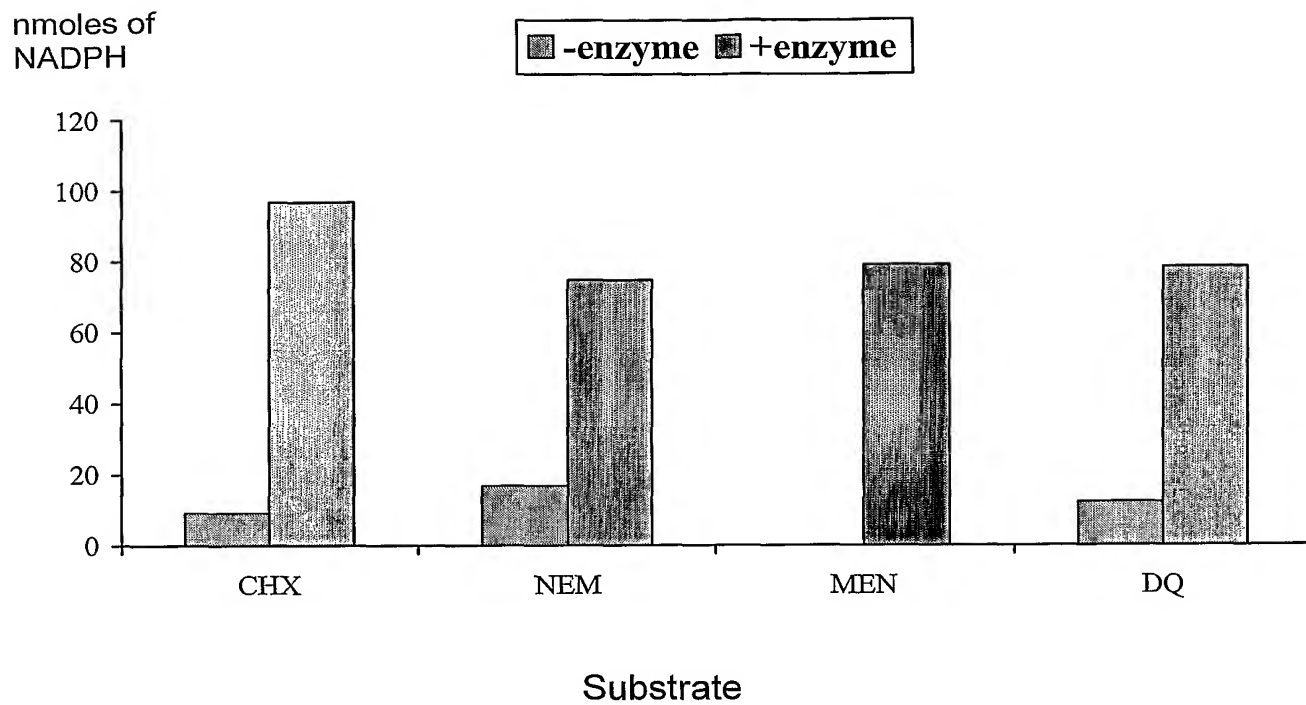


Figure 5: NADPH dehydrogenase activity of recombinant 2031 OR with cyclohexenone (CHX), N-ethylmaleimide (NEM), menadione (MEN) or duroquinone (DQ) as substrates. Final concentrations in the assay were as follows: 500 μ M substrate, 120 μ M NADPH, 1 μ g/200 μ L 2031 OR.

17/17

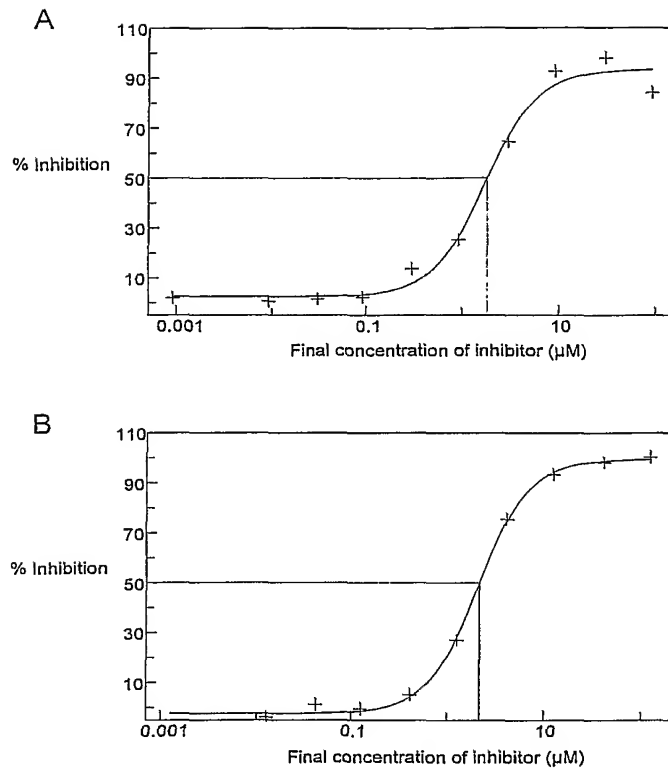


Figure 6: Inhibition of 2031 OR function by two inhibitors (shown in A and B) identified by high-throughput screening.